Tail exactness of multivariate saddlepoint approximations

O.E. Barndorff-Nielsen[*]   C. Klüppelberg[†]

Abstract

We consider a log-concave density f in $\mathbb{R}^m$ satisfying certain weak conditions, particularly on the Hessian matrix of $-\log f$. For such a density, we prove tail exactness of the multivariate saddlepoint approximation. The proof is based on a local limit theorem for the exponential family generated by f. However, the result refers not to asymptotic behaviour under repeated sampling, but to a limiting property at the boundary of the domain of f. Our approach does not apply any complex analysis, but relies totally on convex analysis and exponential models arguments.

Running headline: Multivariate saddlepoint approximations.

AMS 1991 Subject Classifications: primary: 62E17, 62E20, 62F11, 62H10
secondary: 60E10, 60F05

Keywords: asymptotic normality, convex analysis, exponential models, local limit theorem, moment generating function, Legendre transform, saddlepoint approximation.

[*]Department of Mathematical Sciences, Aarhus University, DK-8000 Aarhus C, Denmark, email: atsoebn@mi.aau.dk

[†]Center for Mathematical Sciences, Munich University of Technology, D-80290 Munich, Germany, email: cklu@mathematik.tu-muenchen.de, http://www.mathematik.tu-muenchen.de

1 Introduction

In a wide range of interesting cases saddlepoint approximations of probability distributions turn out to be highly accurate not only in large deviations regions, but even at the very extremes of the distributions. This is a surprising and useful, but poorly understood phenomenon (cf. for instance Barndorff-Nielsen and Cox (1989, 1994), or Whittle (1993): "A conversation with Henry Daniels", see p. 350)). The asymptotic derivations of saddlepoint approximations which refer to increasing sample size give no clues to the question.

The present paper considers the problem from a geometric or convex analysis viewpoint and provides some simple sufficient conditions ensuring that the (unnormalised) saddlepoint approximation is not only accurate but in fact becoming exact as one approaches the boundary of the support of the associated probability distributions. Our discussion concerns multivariate distributions, the much simpler univariate case having been treated in Barndorff-Nielsen and Klüppelberg (1992).

For one-dimensional distributions the tail accuracy of the saddlepoint approximation has been studied in very considerable detail by Daniels (1954) and Jensen (1988, 1989). The above-mentioned paper by Barndorff-Nielsen and Klüppelberg (1992) complements their work, see our discussion given there. Already in the one-dimensional case non-normal limiting behaviour is not uncommon; for a characterisation of all possible limit laws see Balkema, Klüppelberg and Resnick (1997). This phenomenon can be expected to be much more complex in the multivariate case, although we have no results to illustrate that. In fact, our paper is but a small first step towards unravelling the complexity of the multivariate situation.

We consider densities in $\mathbb{R}^m$ of the form

$$f(t) = e^{-\psi(t)} ,$$

where the function ψ is a convex function in C^2, and we show that, under some further regularity conditions, the (unnormalised) saddlepoint approximation $f^\dagger(t)$ of $f(t)$ becomes exact as t approaches the boundary of the domain of ψ. (For general discussions of saddlepoint approximations and their role in statistics, see Barndorff-Nielsen and Cox (1989, 1994) and Jensen (1995).)

The density f generates an exponential family

$$f_\lambda(t) = e^{\langle \lambda, t \rangle - \psi(t)} / C(\lambda)$$

for $\lambda \in \mathbb{R}^m$ such that $C(\lambda) = \int e^{\langle \lambda, t \rangle} f(t) dt < \infty$.

The conditions which we impose on ψ imply that f_λ is asymptotically normal; i.e. there exist norming constants $a_\lambda > 0$ and $b_\lambda \in \mathbb{R}^m$ such that the normalised densities satisfy

$$g_\lambda(t) = a_\lambda f_\lambda(b_\lambda + a_\lambda t) \to \varphi(t)$$

uniformly on $\mathbb{R}^m$ for λ tending to a boundary point of the domain of possible values of λ and where φ is a non-degenerate normal density in $\mathbb{R}^m$. Such results have been derived for a slightly more general class of densities in $\mathbb{R}$ in Balkema, Klüppelberg and Resnick (1993), the latter is henceforth abbreviated as BKR (1993).

The results of BKR (1993) have been applied in Barndorff-Nielsen and Klüppelberg (1992) to show that for a rather wide class of log-concave densities in $\mathbb{R}$ the saddlepoint approximation becomes exact in the tail. In the present paper we generalise this result to a multivariate setting.

The paper is organised as follows. In Section 2 we set the stage and present some analytic concepts. The main result there states uniform asymptotic normality of the exponential family f_λ which yields the asymptotic behaviour of the moment generating function $C(\lambda)$ as λ tends to boundary points of the domain of C. This generalises Theorem 6.6 of BKR (1993). In Section 3 we show, under regularity conditions, that the saddlepoint approximation $f^\dagger(t) \sim f(t)$ as t tends to boundary points of the domain of f. Here $\sim$ means that the quotient of lhs and rhs tends to 1. In Section 4 we introduce conditions, which are easy to check for many examples, where our results hold. Section 5 concludes the paper with some examples.

2 Stage setting

We consider an m-dimensional probability distribution having density with respect to Lebesgue measure of the form

$$f(t) = e^{-\psi(t)}, \quad t \in \mathbb{R}^m, \tag{2.1}$$

where ψ is a closed convex function on $\mathbb{R}^m$. The domain of ψ is denoted by $\text{dom}\psi$, i.e.

$$\text{dom}\psi = \{t \in \mathbb{R}^m : \psi(t) < \infty\},$$

and $D = \text{int}(\text{dom}\psi)$ indicates the interior of the domain of ψ. We use definitions and properties of convex analysis and exponential models in accordance with Rockafeller (1970)
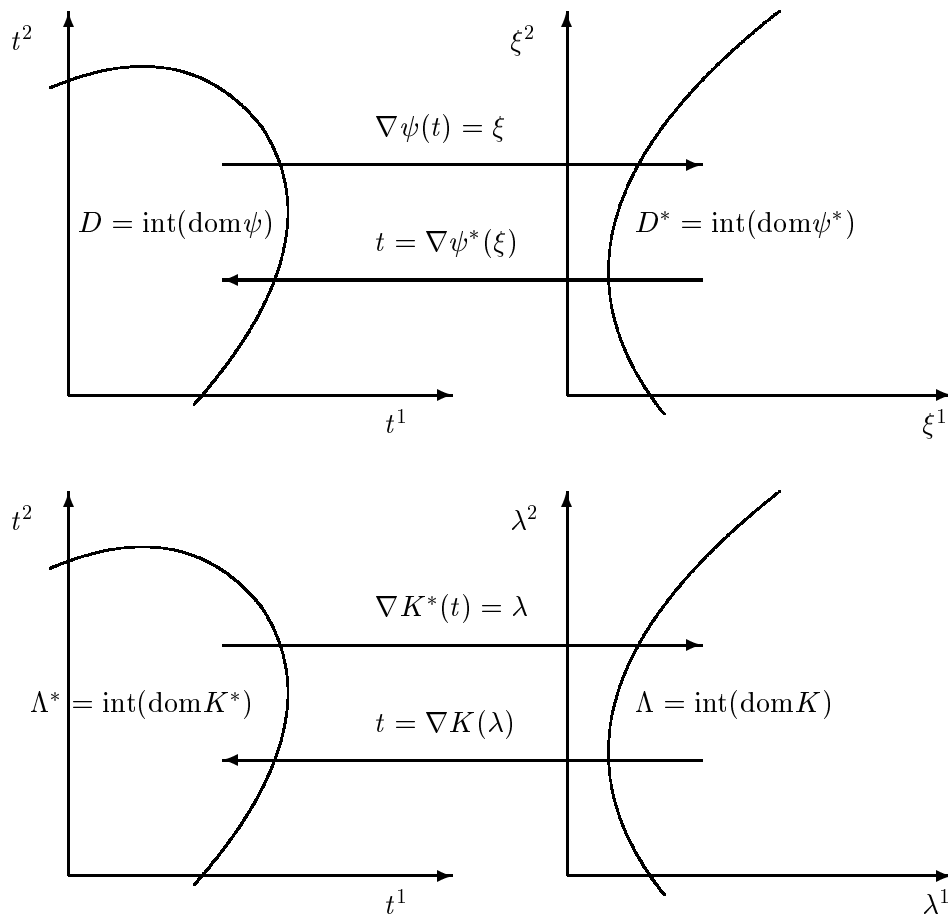
Figure 2.1 *Interiors of domains of $\psi(t) = -\ln f(t)$ and $K(\lambda) = \ln C(\lambda)$ and their convex conjugates (two-dimensional indication), see conditions (iii)-(v). Notice that we can identify D and Λ^* by condition (iv) and D^* and Λ by condition (v).*

and Barndorff-Nielsen (1978), the latter henceforth abbreviated BN (1978). The conjugate ψ^* of ψ is defined as the function on $\mathbb{R}^m$ with values

$$\psi^*(\xi) = \sup_{t \in \mathbb{R}^m} \{\langle \xi, t \rangle - \psi(t)\}, \quad \xi \in \mathbb{R}^m. \tag{2.2}$$

This conjugate function is also closed and convex, and the conjugate of ψ^* equals ψ, i.e. $\psi^{**} = \psi$. Let $D^* = \mathrm{int}(\mathrm{dom}\psi^*)$.

The exponential family $\mathcal{F}$ generated by the density f and the identity mapping on $\mathbb{R}^m$ consists of all probability measures P_λ with density of the form

$$f_\lambda(t) = e^{\langle \lambda, t \rangle - \psi(t)} / C(\lambda) \tag{2.3}$$

for $\lambda \in \mathbb{R}^m$ such that $C(\lambda) = \int e^{\langle \lambda, t \rangle} f(t) dt < \infty$. The cumulant function $K(\lambda) = \ln C(\lambda)$

4

is a closed convex function on $\mathrm{dom}K = \{\lambda \in \mathbb{R}^m \ : \ K(\lambda) < \infty\}$, and the conjugate of K is denoted by K^*. Let

$$\Lambda = \mathrm{int}(\mathrm{dom}K) \quad \text{and} \quad \Lambda^* = \mathrm{int}(\mathrm{dom}K^*).$$

We use the following notation: $\nabla\psi$ and $\nabla^2\psi$ denote the gradient and the Hessian of ψ respectively. Let Σ be an arbitrary square root of $(\nabla^2\psi)^{-1}$. Even though we have not assumed Σ symmetric we use Σ^{-2} as shorthand for $(\Sigma\Sigma^T)^{-1}$.

Throughout the paper we will work under the following conditions (where (i) was introduced above).

(i) ψ is convex and closed.

(ii) $\psi \in C^2$ and $\nabla^2\psi$ is positive definite on D.

(iii) $\nabla\psi$ is a one–to–one mapping from D onto D^*, with inverse $\nabla\psi^*$, and if $t \in D$ and $\xi = \nabla\psi(t)$ then

$$\langle t, \xi \rangle = \psi(t) + \psi^*(\xi).$$

(iv) $\mathrm{int}(\mathrm{dom}\psi) = \mathrm{int}(\mathrm{dom}K^*)$

(v) $\mathrm{int}(\mathrm{dom}K) = \mathrm{int}(\mathrm{dom}\psi^*)$

(vi) Let Σ be an arbitrary square root of $(\nabla^2\psi)^{-1}$ and t_n denote a fixed, but arbitrary sequence of points in D. We assume that as $n \to \infty$

$$\Sigma^T(t_n)\Sigma^{-2}(t_n + x\Sigma(t_n))\Sigma(t_n) \to I \quad \text{locally uniformly in } x \in \mathbb{R}^m. \tag{2.4}$$

The matrix I denotes the identity matrix and by local uniform convergence we mean that the limit relation in (2.4) holds uniformly in x on each compact subset of $\mathbb{R}^m$. In particular, this requires that for all $K_0 \subset \mathbb{R}^m$, K_0 compact, we have that $t_n + K_0\Sigma(t_n) \subset D$ for all $n \geq n(K_0)$.

The conditions (iii)–(v) hold under a fairly simple condition (in addition to (i) and (ii)), as shown in Theorem 4.1 below.

We draw attention in particular to the properties (iv) and (v) which are illustrated in Figure 2.1.

5

In general we think of vectors as row vectors. In certain formulas we shall have to work with column vectors; the appropriate interpretation will, however, be clear from the operation at hand. For notational ease we shall therefore refrain from indicating transposition of vectors but rely on the common sense of the reader.

Definition 2.2 A matrix function Σ which satisfies (2.4) is called *self-neglecting* with respect to the sequence t_n.

Suppose Σ is self-neglecting. For $t \in D$ define a function φ_t with argument $x \in \mathbb{R}^m$ by

$$\varphi_t(x) = \psi(t + x\Sigma(t)) - \psi(t) - x\Sigma(t)\nabla\psi(t). \tag{2.5}$$

Then for any given $M > 0$ and $\epsilon > 0$

$$|\varphi_{t_n}(x) - \frac{1}{2}\langle x, x\rangle| \le \frac{\epsilon}{2}\langle x, x\rangle, \quad \|x\| \le M,$$

holds for all sufficiently large n. This follows simply via Taylor expansion invoking (2.4).

By this property we can derive convergence of the properly standardised exponential family to a parabola on compact sets. For our purpose, however, this does not suffice, we need to control the tails as well. As in the one-dimensional case, we use a concept, called ANET, which has been introduced in BKR (1993, Section 6). We summarise the definition and some properties of ANET from that paper.

Definition 2.3 (a) A sequence of random vectors (X_n) in $\mathbb{R}^m$ is *asymptotically normal* if there exist positive affine transformations A_n on $\mathbb{R}^m$ such that $A_n(X_n) \overset{d}{\to} N$ as $n \to \infty$, where N is a standard normal vector.

(b) The sequence (X_n) is *asymptotically normal with exponential tails (ANET)* if the vectors $A_n(X_n)$ have densities g_n which satisfy the condition: For any $\epsilon > 0$ there exists an index n_0 such that for $n \ge n_0$

$$|g_n(t) - \varphi(t)| < \epsilon e^{-\|t\|/\epsilon}, \quad t \in \mathbb{R}^m, \tag{2.6}$$

where φ is the m-dimensional standard normal density. □

The reason for the name "with exponential tails" comes from the fact that the tails of $g_n(t)$ (for $t \to \infty$) decrease exponentially fast eventually (see Proposition 6.1 of BKR (1993)).

We want to apply this concept in the following situation:

Let $X = (X_1, \ldots, X_m)$ be a random vector with density $f = e^{-\psi}$ such that conditions (i) – (vi) are satisfied. Denote further by $X_\lambda = (X_{1,\lambda}, \ldots, X_{m,\lambda})$ a random vector distributed according to the exponential family density f_λ in (2.3). Now we consider $t \in D$ as a function of λ by

$$t = t(\lambda) = \nabla \psi^*(\lambda), \tag{2.7}$$

or, analogously, we consider $\lambda \in \Lambda$ as a function of t by

$$\lambda = \lambda(t) = \nabla \psi(t). \tag{2.8}$$

The following result generalises Theorem 6.6 of BKR (1993).

Theorem 2.4 *Let $X = (X_1, \ldots, X_m)$ have density $f = e^{-\psi}$ such that conditions (i)–(vi) hold. Define the exponential family f_λ as in (2.3). Let t_n be a sequence of points in D such that Σ is self–neglecting with respect to t_n, let $\lambda_n = \nabla \psi(t_n)$ and suppose that $\lambda_n \in \Lambda$. Then the normalised densities*

$$g_{\lambda_n}(x) = |\Sigma(t_n)| f_{\lambda_n}(t_n + x\Sigma(t_n))$$

of $(X_{\lambda_n} - t_n)\Sigma^{-1}(t_n)$ are ANET as $n \to \infty$. Furthermore, the moment generating function $C(\lambda)$ of f and the conjugate transform ψ^* of ψ are related by the asymptotic equality

$$C(\lambda_n) \sim (2\pi)^{m/2} |\Sigma(t_n)| e^{\psi^*(\lambda_n)}, \quad n \to \infty. \tag{2.9}$$

PROOF. For $t \in D$, let

$$g_\lambda(x) = \frac{|\Sigma(t)|}{C(\lambda)} \exp\{\langle t + x\Sigma(t), \lambda\rangle - \psi(t + x\Sigma(t))\}.$$

By Taylor expansion we have (for some $|\theta| \leq 1$):

$$\psi(t + x\Sigma(t)) = \psi(t) + x\Sigma(t)\nabla\psi(t) + \frac{1}{2}x\Sigma(t)^T \nabla^2 \psi(t + \theta x \Sigma(t))\Sigma(t)x.$$

Inserting this in the expression for $g_\lambda(x)$ and taking $\lambda = \nabla\psi(t)$ and using condition (iii) we obtain

$$g_\lambda(x) = \frac{|\Sigma(t)|}{C(\lambda)} \exp\left\{\psi^*(\lambda) - \frac{1}{2}x\Sigma^T(t)\nabla^2 \psi(t + \theta x\Sigma(t))\Sigma(t)x\right\}.$$

7

Now, inserting t_n for t and λ_n for λ we find, by the self–neglecting property (2.4), that

$$g_{\lambda_n}(x) = \frac{|\Sigma(t_n)|}{C(\lambda_n)} \exp\{\psi^*(\lambda_n) - \frac{1}{2}\langle x, x \rangle + r_{\lambda_n}(x)\}$$

where $r_{\lambda_n}(x) \to 0$ as $n \to \infty$ locally uniformly in x. Hence for $n \to \infty$,

$$g_{\lambda_n}(x) \sim d(\lambda_n) e^{-\frac{1}{2}\langle x, x \rangle}$$

locally uniformly in x, where

$$d(\lambda_n) = \frac{|\Sigma(t_n)|}{C(\lambda_n)} e^{\psi^*(\lambda_n)}.$$

Now since $\psi = -\log f$ is convex, also $-\log g_\lambda$ is convex and by Proposition 6.5 of BKR (1993) g_{λ_n} is ANET for $n \to \infty$. In particular, $d(\lambda_n) \to (2\pi)^{-m/2}$ which gives (2.9). $\qquad \square$

From the Corollary to Proposition 6.3 of BKR (1993) we obtain immediately

Corollary 2.5 *In the setting of Theorem 2.4 we have*

$$(X_{\lambda_n} - t_n)\Sigma^{-1}(t_n) \overset{d}{\to} N(0, I).$$

Moreover,

$$(E(X_{\lambda_n}) - t_n)\Sigma^{-1}(t_n) \to 0 \quad and \quad \Sigma^{-1}(t_n)V_{\lambda_n}\Sigma^{-1}(t_n) \to I,$$

where V_λ is the covariance matrix of the vector X_λ. $\qquad \square$

3 Tail exactness of the saddlepoint approximation

We now use the results of Section 2 to prove tail exactness of the multivariate saddlepoint approximation:

Theorem 3.1 *Let $f = e^{-\psi}$ be a density satisfying properties* (i) − (vi) *for a sequence t_n and define $\lambda_n = \nabla\psi(t_n)$. Then the saddlepoint approximation of $f(t_n)$ given by*

$$f^\dagger(t_n) = (2\pi)^{-m/2}|\nabla^2\psi(t_n)|^{1/2} \exp\{K(\lambda_n) - \langle \lambda_n, t_n \rangle\}$$

becomes asymptotically exact; i.e.

$$\lim_{n \to \infty} f^\dagger(t_n)/f(t_n) = 1.$$

8

Remark The usual form of the saddlepoint approximation would have $|\nabla^2 K(\lambda_n)|^{-1/2}$ instead of $|\nabla^2 \psi(t_n)|^{1/2}$ in the expression for $f^\dagger$. However, as will be shown later, the two expressions are asymptotically equivalent under the conditions given in Section 4.

PROOF. The density f satisfies the conditions of Theorem 2.4 with scale matrix Σ; i.e. (2.9) holds and hence for $n \to \infty$

$$K(\lambda_n) = \log C(\lambda_n) = \psi^*(\lambda_n) + \log\{(2\pi)^{m/2}|\Sigma(t_n)|\} + o(1) ,$$

where $t_n = \nabla \psi^*(\lambda_n)$. As before we write $|\Sigma(t)|$ for $|\Sigma(t)\Sigma^T(t)|^{1/2}$. This implies

$$
\begin{aligned}
K(\lambda_n) - \langle \lambda_n, t_n \rangle &= \psi^*(\lambda_n) + \log\{(2\pi)^{m/2}|\Sigma(t_n)|)\} - \langle \lambda_n, t_n \rangle + o(1) \\
&= -\psi(t_n) + \log\{(2\pi)^{m/2}|\Sigma(t_n)|\} + o(1) \\
&= \log\{f(t_n)(2\pi)^{m/2}|\Sigma(t_n)|\} + o(1).
\end{aligned}
$$

Hence

$$\exp\{K(\lambda_n) - \langle \lambda_n, t_n \rangle\} \sim (2\pi)^{m/2}|\Sigma(t_n)|f(t_n), \quad n \to \infty ,$$

and, since $\Sigma(t_n)^{-2} = \nabla^2 \psi(t_n)$, the result follows. $\square$

4   A simple sufficient condition

The following theorem shows, in particular, that under conditions (i) and (ii) the properties (iii)–(v) assumed in Section 2 hold in a simple setting.

Theorem 4.1 *Suppose that ψ is C^2 with positive definite Hessian $\nabla^2 \psi$ on D, and that ψ is steep, i.e.*

$$\frac{d}{d\lambda}\psi(t_0 + \lambda(t - t_0)) \downarrow -\infty , \quad \lambda \downarrow 0 ,$$

for any $t \in D$ and any point t_0 of the boundary of D. Then

(a) *The gradient $\nabla \psi$ is a one-to-one function from $D = \mathrm{int}(\mathrm{dom}\psi)$ onto $D^* = \mathrm{int}(\mathrm{dom}\psi^*)$, with inverse $\nabla \psi^*$.*

(b) *For $\xi \in D^*$ we have*

$$\psi^*(\xi) = \langle \xi, t \rangle - \psi(t) , \tag{4.1}$$

where t is the unique solution of

$$\nabla \psi(t) = \xi . \tag{4.2}$$

(c) *The domain of the cumulant function K is open and satisfies*

$$\Lambda = \mathrm{dom}\, K = \mathrm{int}(\mathrm{dom}\,\psi^*) = D^*$$

(In particular, then, the exponential family $\mathcal{F}$ is regular.)

(d) *The conjugate K^* has open domain*

$$\Lambda^* = \mathrm{dom}\, K^* = \mathrm{int}(\mathrm{dom}\,\psi) = D \,, \tag{4.3}$$

and for any $t \in D$ we have

$$K^*(t) = \langle \lambda, t \rangle - K(\lambda) \,, \tag{4.4}$$

where λ is the unique solution of the equation

$$\nabla K(\lambda) = t \,. \tag{4.5}$$

PROOF. The assumptions of the theorem imply that ψ is essentially smooth (i.e. $D \neq \emptyset$, ψ is differentiable on D and steep) and essentially strictly convex (BN (1978), Theorems 5.28 and 5.29). This, in turn, has the consequence (Theorem 5.30, loc. sit.) that the conjugate ψ^* of ψ is also essentially smooth and essentially strictly convex. Conclusions (a) and (b) now follow from Theorem 5.33 in BN (1978).

Next we consider the cumulant function K. Writing $\psi_\lambda(x) = \psi(x) - \langle \lambda, x \rangle$, we find on account of Theorem 6.1 in BN (1978) that

$$K(\lambda) < \infty \Leftrightarrow 0 \in \mathrm{int}(\mathrm{dom}\,\psi_\lambda^*) \Leftrightarrow \lambda \in \mathrm{int}(\mathrm{dom}\,\psi^*) \tag{4.6}$$

verifying (c).

The regularity of the exponential family $\mathcal{F}$ implies (BN (1978), Theorem 9.2) that (4.3) holds, and Theorems 9.1 and 9.13 (loc. cit.) then show the validity of (4.4) and (4.5).

$\square$

Remark (i) A convex function ψ which is differentiable on D is steep if $\|\nabla\psi(t_n)\| \to \infty$ for all sequences $(t_n)_{n\in\mathbb{N}}$ in D converging to t_0 for all boundary points t_0 of D.

(ii) If ψ is a closed convex and differentiable function such that $D = \mathrm{dom}\,\psi$, i.e. the domain of ψ is open, then ψ is steep (cf. Rockafellar (1970)).

10

Corollary 4.2 *Under the conditions of Theorem 4.1 we have*

$$\langle \lambda_n, t_n \rangle = K(\lambda_n) + K^*(t_n),$$

$$\nabla^2 K(\lambda_n) \cdot \nabla^2 K^*(t_n) = I,$$

$$\nabla^2 K(\lambda) = V_\lambda$$

for $\lambda \in D^*$, V_λ denoting the variance of X_λ. In this case $f^\dagger$ may (in view of Corollary 2.5) be rewritten as

$$f^\dagger(t_n) = (2\pi)^{-m/2} |\nabla^2 K(\lambda_n)|^{-1/2} \exp\{K(\lambda_n) - \langle \lambda_n, t_n \rangle\}$$

which is the classical saddlepoint approximation of $f(t_n)$. Alternatively, $f^\dagger(t_n)$ may be expressed as

$$f^\dagger(t_n) = (2\pi)^{-m/2} |\nabla^2 K^*(t_n)|^{1/2} \exp\{-K^*(t_n)\}. \qquad \square$$

5 Examples

One–dimensional densities $f(t) = e^{-\psi(t)}$ where ψ is C^2 with $\psi'' > 0$ satisfying (2.4) have been considered in BKR (1993) and Barndorff–Nielsen and Klüppelberg (1992). Examples of possible functions ψ are

$$t^\alpha, \quad \alpha > 1$$

$$e^t$$

$$t - t^\alpha, \quad 0 < \alpha < 1$$

$$t \log t$$

$$(-t)^\alpha, \quad \alpha > 0$$

In the first four examples we consider $t \uparrow t_0 = \infty$, in the last one $t \uparrow t_0 = 0$. The self–neglecting property (2.4) reduces to

$$\lim_{t \uparrow t_0} \frac{\sigma(t + x\sigma(t))}{\sigma(t)} = 1 \quad \text{locally uniformly in } x \in \mathbb{R}, \tag{5.1}$$

where $\sigma = 1/\sqrt{\psi''}$. If $t_0 = \infty$, then sufficient for σ to be self–neglecting is that σ has a derivative σ' satisfying $\sigma'(t) = 0$ as $t \to \infty$. If $t_0 < \infty$ and both σ and σ' vanish at t_0, then σ is self–neglecting (Resnick (1987), Lemmas 1.2 and 1.3).

The major problem for explicit examples in the multivariate case will be to check the self-neglecting condition (2.4). In some cases asymptotic estimates are possible; in some cases Σ can be calculated explicitly.

In particular, for a 2×2 non-singular symmetric matrix

$$\Sigma^{-2} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

the eigenvalues ρ_1, ρ_2 are

$$\rho_{1,2} = \frac{1}{2}(a + c \pm \sqrt{(a-c)^2 + 4b^2})$$

and the eigenvectors are

$$e_i = \begin{pmatrix} b \\ a - \rho_i \end{pmatrix}, \quad i = 1, 2.$$

Then the symmetric square root of Σ^2, which we now write as

$$\Sigma = \begin{pmatrix} r & s \\ s & v \end{pmatrix}$$

satisfies the equation

$$\begin{pmatrix} r & s \\ s & v \end{pmatrix} \begin{pmatrix} b \\ a - \rho_i \end{pmatrix} = \frac{1}{\sqrt{\rho_i}} \begin{pmatrix} b \\ a - \rho_i \end{pmatrix}.$$

Hence, for $\rho_1 \neq \rho_2$ we obtain

$$\begin{aligned}
v &= \frac{1}{\rho_1 - \rho_2} \left(\frac{a - \rho_1}{\sqrt{\rho_1}} - \frac{a - \rho_2}{\sqrt{\rho_2}} \right) \\
s &= \frac{b}{\rho_1 - \rho_2} \left(\frac{1}{\sqrt{\rho_1}} - \frac{1}{\sqrt{\rho_2}} \right) \\
r &= \frac{1}{\sqrt{\rho_1}} - \frac{\rho_1 - a}{\rho_1 - \rho_2} \left(\frac{1}{\sqrt{\rho_1}} - \frac{1}{\sqrt{\rho_2}} \right)
\end{aligned}$$

In the following situation the problem can be reduced to an essentially one–dimensional equivalent. Assume that ψ is rotationally invariant, i.e.

$$\psi(t) = g(r), \quad \text{where} \quad r = \|t\|.$$

Write $t = (t^1, t^2) = (r \cos \theta, r \sin \theta)$ in polar coordinates. Then by straightforward analysis, for $\psi(t^1, t^2) = g(r, \theta)$,

$$\Sigma^{-2}(t^1, t^2) = \nabla^2 \psi(t^1, t^2) = \begin{pmatrix} \cos^2 \theta g''(r) + \sin^2 \theta \frac{1}{r} g'(r) & \sin \theta \cos \theta (g''(r) - \frac{1}{r} g'(r)) \\ \sin \theta \cos \theta (g''(r) - \frac{1}{r} g'(r)) & \sin^2 \theta g''(r) + \cos^2 \theta \frac{1}{r} g'(r) \end{pmatrix}.$$

12

Since ψ is rotationally invariant, it suffices to consider $\theta = 0$, hence for $(t^1, t^2) = (r, 0)$ in polar coordinates the matrix $\Sigma^{-2}(t^1, t^2)$ reduces to

$$\Sigma^{-2}(r) = \begin{pmatrix} g''(r) & 0 \\ 0 & g'(r)/r \end{pmatrix}$$

which is positive definite provided $g''(r) > 0$ and $g'(r)/r > 0$, and then ψ is convex. The self-neglecting condition (2.4) reduces then to

$$\Sigma(r)\Sigma^{-2}((r, 0) + x\Sigma(r))\Sigma^T(r) \to 1 \tag{5.2}$$

as r tends to ∞.

Proposition 5.1 *Let $f = e^{-\psi}$ be a density on $\mathbb{R}^2$ such that ψ is rotationally invariant. Write $\psi(t) = g(r)$ for $r = \|t\|$. Assume that g is C^2 and satisfies the following conditions as $r \to \infty$:*

(a) $g'(r)/r \to \infty$ and $g''(r) \to \infty$,

(b) $0 < \liminf \dfrac{g'(r)}{rg''(r)} \leq \limsup \dfrac{g'(r)}{rg''(r)} < \infty$,

(c) the function $1/\sqrt{g''}$ is self-neglecting in the sense of (5.1),

(d) $\dfrac{g'(r + x/\sqrt{g''(r)})}{g'(r)} \to 1$ locally uniformly in x.

Then Theorems 2.4 and 3.1 hold and the saddlepoint approximation $f^{\dagger}$ of f becomes exact as $r \to \infty$.

PROOF. By condition (a) above and Remark (i) after Theorem 4.1 we work in the framework of Theorem 4.1. Hence we only have to check the self-neglecting condition (5.2). Write $x = (x^1, x^2)$, then

$$(r, 0) + x\Sigma(r) = \left(r + \frac{x^1}{\sqrt{g''(r)}}, \frac{x^2}{\sqrt{g''(r)}} \right) = (\widetilde{r}, \widetilde{\theta})$$

giving

$$\Sigma^T(r)\Sigma^{-2}(\widetilde{r}, \widetilde{\theta})\Sigma(r) = \begin{pmatrix} \cos^2\widetilde{\theta}\dfrac{g''(\widetilde{r})}{g''(r)} + \sin^2\widetilde{\theta}\dfrac{g'(\widetilde{r})/\widetilde{r}}{g''(r)} & \sin\widetilde{\theta}\cos\widetilde{\theta}\dfrac{g''(\widetilde{r}) - g'(\widetilde{r})/\widetilde{r}}{\sqrt{g''(r)g'(r)/r}} \\ \sin\widetilde{\theta}\cos\widetilde{\theta}\dfrac{g''(\widetilde{r}) - g'(\widetilde{r})/\widetilde{r}}{\sqrt{g''(r)g'(r)/r}} & \sin^2\widetilde{\theta}\dfrac{g''(\widetilde{r})}{g'(r)/r} + \cos^2\widetilde{\theta}\dfrac{g'(\widetilde{r})/\widetilde{r}}{g'(r)/r} \end{pmatrix} .$$

First notice that $\widetilde{\theta} = x^2/\sqrt{g''(r)} \to 0$ by (a). Conditions (c) and (d) guarantee that locally uniformly

$$g''(\widetilde{r}) \sim g''(r) \quad \text{and} \quad g'(\widetilde{r})/\widetilde{r} \sim g'(r)/r \,.$$

Hence the quotient $\dfrac{g'(\widetilde{r})/\widetilde{r}}{g''(r)}$ remains bounded away from 0 and ∞. Furthermore,

$$\frac{g''(\widetilde{r}) - g'(\widetilde{r})/\widetilde{r}}{\sqrt{g''(r)g'(r)/r}} \quad \sim \quad \frac{g''(\widetilde{r}) - g'(\widetilde{r})/\widetilde{r}}{\sqrt{g''(\widetilde{r})g'(\widetilde{r})/\widetilde{r}}} \quad = \quad \sqrt{\frac{g''(\widetilde{r})}{g'(\widetilde{r})/\widetilde{r}}} - \sqrt{\frac{g'(\widetilde{r})/\widetilde{r}}{g''(\widetilde{r})}} \,,$$

where the rhs is also bounded away from 0 and ∞. Hence

$$\Sigma(r)\Sigma^{-2}(\widetilde{r}, \widetilde{\theta})\Sigma^T(r) \to I \,.$$

This concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Example 5.2 Let

$$\psi(t^1, t^2) = \frac{1}{4}((t^1)^2 + (t^2)^2)^2 \,.$$

Hence in polar coordinates

$$\psi(t^1, t^2) = g(r) = \frac{1}{4}r^4 \,.$$

We check the conditions of Proposition 5.1. Conditions (a) and (b) are immediate. The derivative of the function $1/\sqrt{g''(r)} = 1/(\sqrt{3}r)$ tends to 0 as $r \to \infty$, hence $1/\sqrt{g''(r)}$ is self-neglecting. Furthermore,

$$\frac{g'(r + x/(\sqrt{3}r))}{g'(r)} = \left(\frac{r + x/(\sqrt{3}r)}{r}\right)^3 \to 1 \,. \qquad\qquad\qquad\qquad \square$$

The following describes a more general set-up, which covers for instance the m-dimensional version of Example 5.2 for $m \geq 2$. Let $f = e^{-\psi}$ be a density on $\mathbb{R}^m$, where ψ is of the form

$$\psi(t) = h(g(t^1) + \cdots + g(t^m)) \,,$$

where $g, h \in C^2$. Denote

$$h' = h'(g(t^1) + \cdots + g(t^m)) \quad \text{and} \quad h'' = h''(g(t^1) + \cdots + g(t^m)) \,.$$

Let furthermore

$$\begin{aligned}
\varphi &= (g'(t^1), \ldots, g'(t^m)) \\
\overline{\varphi} &= \left(\frac{g'(t^1)}{\sqrt{g''(t^1)}}, \ldots, \frac{g'(t^m)}{\sqrt{g''(t^m)}}\right) \,,
\end{aligned}$$

14

$$v = \sqrt{h''/h'}\,\overline{\varphi}\,,$$

$$a = (1 + \langle v, v \rangle)^{-1} = \left(1 + \frac{h''}{h'} \sum_{i=1}^{m} \frac{g'(t^i)^2}{g''(t^i)}\right)^{-1},$$

$$b = a/(1 + \sqrt{a})\,.$$

Proposition 5.3 *Assume that in the above setting the following conditions hold.*

(a) $h'(t) > 0, h''(t) \geq 0$ *for* $t = g(t^1) + \cdots + g(t^m)$ *and* $g''(t^i) > 0$ *for* $i = 1, \ldots, m$.

(b) $h'(t_n + x\Sigma(t_n))/h'(t_n) \to 1$ *locally uniformly as* $n \to \infty$.

(c) *Denote by* $\delta(t)$ *the diagonal matrix with elements* $g''(t^i)$, $i = 1, \ldots, m$, *then for* $v = v(t_n)$, $b = b(t_n)$, $\delta = \delta(t_n)$ *and* $\widetilde{\delta} = \delta(t_n + x\Sigma(t_n))$,

$$[I - bv^T v]\widetilde{\delta}\delta^{-1}[I - bv^T v]^{-1} \to I\,,$$

locally uniformly as $n \to \infty$.

(d) *For* $v = v(t_n)$, $b = b(t_n)$ *and* $\widetilde{v} = v(t_n + x\Sigma(t_n))$, *we have*

$$[I - b\widetilde{v}^T \widetilde{v}]^{-1}[I - bv^T v] \to I\,,$$

locally uniformly as $n \to \infty$.

Then Theorems 2.4 and 3.1 hold and the saddlepoint approximation $f^\dagger$ *of* f *becomes exact as* $n \to \infty$.

PROOF. Notice first that with the above notation $\nabla\psi(t) = h'\varphi$ and

$$\nabla^2\psi(t) = h''\varphi^T\varphi + h'\delta\,. \tag{5.3}$$

Hence for any vectors $\xi, t \in \mathbb{R}^m$ with $\xi \neq 0$ we have

$$\xi\nabla^2\psi(t)\xi = h''\langle\xi,\varphi\rangle^2 + h'\sum_{i=1}^{m}\xi_i^2 g''(t^i) > 0\,.$$

Hence ψ is convex on $\mathbb{R}^m$. Notice further that by (5.3)

$$\Sigma^{-2} = \nabla^2\psi = h'^{1/2}\delta^{1/2}\left[I + \frac{h''}{h'}\overline{\varphi}^T\overline{\varphi}\right]\delta^{1/2}h'^{1/2} = h'^{1/2}\delta^{1/2}[I + v^T v]\delta^{1/2}h'^{1/2}\,.$$

Now since

$$[I + v^T v]^{-1} = I - (1 + \langle v, v \rangle)^{-1}v^T v = I - av^T v\,,$$

15

we have

$$\Sigma^2 = (\nabla^2 \psi)^{-1} = h'^{-1/2} \delta^{-1/2} [I - av^T v] \delta^{-1/2} h'^{-1/2} \,.$$

From this representation we obtain immediately the (non-symmetric) square root

$$\Sigma = h'^{-1/2} \delta^{-1/2} [I - av^T v]^{1/2} = h'^{-1/2} \delta^{-1/2} [I - bv^T v] \,,$$

where $b = a/(1 + \sqrt{a})$. It remains to check the self-neglecting condition (2.4). As before we drop arguments and write Σ for $\Sigma(t)$ and $\widetilde{\Sigma}$ for $\Sigma(t + x\Sigma(t))$ with obvious generalisations of this notation to other functions.

$$
\begin{aligned}
\Sigma^T &\widetilde{\Sigma}^{-2} \Sigma \\
&= \left[h'^{-1/2} \delta^{-1/2} [I - bv^T v] \right]^T \widetilde{h}'^{1/2} \widetilde{\delta}^{1/2} [I + \widetilde{v}^T \widetilde{v}] \widetilde{\delta}^{1/2} \widetilde{h}'^{1/2} \left[h'^{-1/2} \delta^{-1/2} [I - bv^T v] \right] \\
&= \frac{\widetilde{h}'}{h'} [I - bv^T v] \delta^{-1/2} \widetilde{\delta}^{1/2} [I - \widetilde{a}\widetilde{v}^T \widetilde{v}]^{-1} \widetilde{\delta}^{1/2} \delta^{-1/2} [I - bv^T v] \\
&= (1 + o(1)) [I - bv^T v] \delta^{-1/2} \widetilde{\delta}^{1/2} [I - \widetilde{b}\widetilde{v}^T \widetilde{v}]^{-1} [I - \widetilde{b}\widetilde{v}^T \widetilde{v}]^{-1} \widetilde{\delta}^{1/2} \delta^{-1/2} [I - bv^T v] \,.
\end{aligned}
$$

The result now follows from the self-neglecting conditions (b)-(d). $\square$

The simplest example falling into the realm of Proposition 5.3 is the standard normal distribution, where $h(t) = t$, $g(t) = t^2/2$, $\varphi = \overline{\varphi} = (t^1, \dots, t^m)$, $v = 0$, $a = 1$ and $b = 1/2$. In that case, of course, the conclusion of Proposition 5.1 is trivial. Example 5.2 for $m \geq 2$ is, as already indicated, another special case.

6   Acknowledgement

References

[1] Balkema, A.A., Klüppelberg, C. and Resnick, S.I. (1993)
Densities with Gaussian tails.
Proc. London Math. Soc. **66**, 568-588.

[2] Balkema, A.A., Klüppelberg, C. and Resnick, S.I. (1997)
Limit distributions for exponential families.
Submitted to *Bernoulli*.

[3] Barndorff-Nielsen, O.E. (1978)

  Information and exponential families. Wiley, Chichester.

[4] Barndorff-Nielsen, O.E. and Cox, D.R. (1989)

  Asymptotic techniques for use in statistics.

  Chapman and Hall, London.

[5] Barndorff-Nielsen, O.E. and Cox, D.R. (1994)

  Inference and asymptotics.

  Chapman and Hall, London.

[6] Barndorff-Nielsen, O.E. and Klüppelberg, C. (1992)

  A note on the tail accuracy of the univariate saddlepoint approximation.

  Annales de Toulouse (Série 6) **1**, 5-14.

[7] Daniels, H.E. (1954)

  Saddlepoint approximation in statistics. *Ann. Math. Statist.* **25**, 631-650.

[8] Jensen, J.L. (1988)

  Uniform saddlepoint approximations. *Adv. Appl. Prob.* **20**, 622-634.

[9] Jensen, J.L. (1989)

  Uniform saddlepoint approximations and log-concave densities.

  J. Roy. Statist. Soc. B **53**, 157-172.

[10] Jensen, J.L. (1995)

  Saddlepoint approximations. Oxford University Press.

[11] Resnick, S.I. (1987)

  Extreme values, regular variation, and point processes.

  Springer, New York.

[12] Rockafellar, R.T. (1970)

  Convex analysis. Princeton University Press.

[13] Whittle, P. (1993)

  A conversation with Henry Daniels. *Statistical Science* **8**, 342-353.

Corresponding author:

Claudia Klüppelberg

Center for Mathematical Sciences

Munich University of Technology

D-80290 Munich, Germany

email: cklu@mathematik.tu-muenchen.de